DIAGNOSTYKA, 2016, Vol. 17, No. 3

ISSN 1641-6414
e-ISSN 2449-5220

A NEW APPROACH TO DIAGNOSTIC SYMPTOM ASSESSMENT BASED ON INFORMATION CONTENT MEASURES

Tomasz GAŁKA
Institute of Power Engineering – Research Institute
8 Mory St, 01-330 Warszawa, Poland
e-mail: tomasz.galka@ien.com.pl

Abstract

A complex diagnostic object typically generates a large number of diagnostic symptoms. For proper lifetime consumption monitoring it is necessary to select those which best represent technical condition deterioration. Such selection may be based on the Singular Value Decomposition method. The paper presents a novel alternative approach, which employs an information content measure. As the end of object life is approached, symptoms are to an increasing extent dominated by deterministic lifetime consumption processes and therefore become more predictable. Thus symptoms with the highest information content decrease rate should be considered most useful. For a proper assessment, however, symptom sensitivity to condition parameters should also be addressed. A new measure referred to as representativeness factor is proposed. Suitability of such approach is demonstrated for the fluid-flow system of a large steam turbine.

Key words: lifetime consumption, diagnostic symptom, information content

NOWA METODA OCENY SYMPTOMÓW DIAGNOSTYCZNYCH OPARTA NA MIARACH ZAWARTOŚCI INFORMACJI

Streszczenie

Złożone obiekty diagnozowania zwykle są źródłem wielu symptomów diagnostycznych. Dla właściwego monitorowania wyczerpywania żywotności należy wybrać te z nich, które najlepiej odwzorowują degradację stanu technicznego. Wybór ten można oprzeć na metodzie rozkładu względem wartości szczególnej. W pracy przedstawiono nowatorskie alternatywne podejście, w którym wykorzystuje się miarę zawartości informacji. Wraz ze zbliżaniem się do końca życia obiektu symptomy są w coraz większym stopniu określone przez deterministyczne procesy ubytku żywotności, a zatem stają się coraz bardziej przewidywalne. Za najbardziej użyteczne należy zatem uznać symptomy z najszybszym spadkiem zawartości informacji. We właściwej ocenie należy jednak uwzględnić także wrażliwość symptomu na parametry stanu. Zaproponowano nową miarę określoną jako współczynnik reprezentatywności. Przydatność takiego podejścia została zademonstrowana na przykładzie układu przepływowego turbiny parowej.

Słowa kluczowe: wyczerpywanie żywotności, symptom diagnostyczny, zawartość informacji

INTRODUCTION

A complex object typically generates a large number of diagnostic symptoms, each of them being related to condition parameters vector by some specific diagnostic relation. Moreover, these symptoms are typically influenced also by control parameters and interference. In lifetime consumption monitoring it is therefore necessary to select those symptoms which best represent object condition deterioration. A possible approach is to employ methods based on Singular Value Decomposition (SVD), originally proposed by Cempel (see e.g. [1]). The author tested the suitability of this approach to condition monitoring of steam turbine fluid-flow systems [2] with encouraging results. An alternative approach can be based on information content measures.

1. INFORMATION CONTENT MEASURES

The concept of employing an information content measure (ICM) for diagnostic symptoms assessment was put forward by the author [3] on the basis of certain consideration concerning the very nature of diagnostic symptoms. The basic relation

$$\mathbf{S}(\theta) = \Phi[\mathbf{X}(\theta), \mathbf{R}(\theta), \mathbf{Z}(\theta)] \quad (1)$$

implies that diagnostic symptoms vector **S** depends on both deterministic (condition parameters vector **X**) and random (vectors of control **R** and interference **Z**) variables; θ denotes time and Φ is the symptom operator. We may therefore treat any $S_i \in \mathbf{S}$ as a random variable with time-dependent parameters.

Any S_i may thus be analyzed in terms of information content. As the end of object lifetime is approached, influence of condition parameters becomes dominant, as both **R** and **Z** are, for a given object, basically characterized by time-independent statistical distributions. This means that the components of **S** become more deterministic, and this is equivalent to information content measure decrease [4]. The rate of this decrease with time thus quantifies the symptom representativeness.

First information content measure was introduced by Shannon [5] and termed Shannon entropy. It is still widely employed. Alternative measures proposed e.g. by Rényi [4] or Tsallis [6] failed to find widespread practical applications, mainly due to problems with physical interpretation of certain factors that appear in relevant equations. Shannon entropy H was originally conceived for verbal communication and is therefore of discrete nature:

$$H = -K\sum_{i=1}^{n} p_i \log_b p_i \ , \tag{2}$$

where p_i denotes the probability of the ith event:

$$\sum_{i=1}^{n} p_i = 1 \tag{3}$$

and K is a constant dependent on the logarithm base b. Typically b = 2, e or 10, which gives h in bits, nats and dits, respectively. Diagnostic symptoms are, however, characterized by continuous distributions. We may therefore employ a measure known as *continuous* or *differential* entropy h, given by [7]

$$h = -K\int_{-\infty}^{\infty} p(S_i)\log_b p(S_i)dS_i \tag{4}$$

where $p(S_i)$ denotes the symptom probability density distribution. As long as we are interested in the shape of the $h(\theta)$ function rather than its absolute value, both K and b are irrelevant; in the following, $b = e$ was assumed. It has to be noted that differential entropy is not a continuous analogue of the Shannon entropy and may assume negative values (although proper interpretation of this occurrence has not yet been given). As it is the shape of the $h(\theta)$ that is of importance, we may employ differential entropy, which simplifies calculations.

Calculations of $h(\theta)$ employ the moving time window procedure. Within each window, a statistical distribution is fitted to measurement data and parameters of this distribution are plotted against time. In order to perform this operation, a symptom probability distribution type has to be assumed. Some considerations concerning this issue may be found in [3]. In particular, Weibull and gamma distributions have been found suitable. It has been shown by the author, however, that distribution type choice is not critical [8]. An example shown in Fig.1 illustrates that Weibull, gamma and normal distributions yield comparable results, although, formally speaking, normal distribution is not supported by basic considerations employing the Energy Processor model. It is therefore justified to use normal distribution, for which h is given by a simple equation [7]:

$$h = \ln\left(\sigma\sqrt{2\pi e}\right) \tag{5}$$

2. PRE-PROCESSING OF MEASUREMENT DATA

Diagnostic symptom time histories are often characterized by a large number of outliers. This is typically the case for complex industrial objects operated in plant environment, with numerous sources of interference.

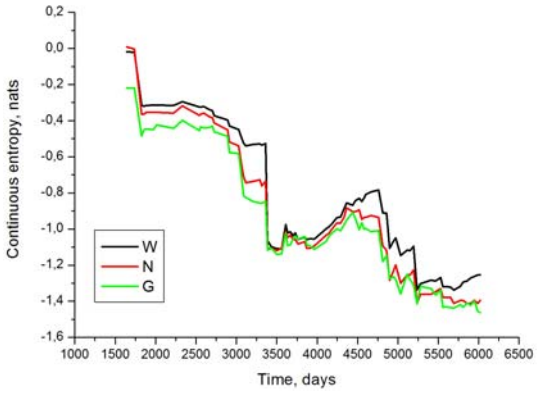

Fig. 1. An example of continuous entropy time histories obtained with Weibull (W), gamma (G) and normal (N) distributions (after [8])

There is no generally accepted and precise definition of an outlier. According to Grubbs [9], 'an outlying observation, or outlier, is one that appears to deviate markedly from other members of the sample in which it occurs'. From the point of view of information theory outliers are equivalent to noise and should be removed. This may be accomplished by so-called peak trimming [3]. This procedure is based on the assumption that, if for the kth symptom value reading $S_i(\theta_k)$

$$S_i(\theta_k)/S_i(\theta_{k-1}) > c_h \ \text{ and } \ S_i(\theta_k)/S_i(\theta_{k+1}) > c_h \tag{6}$$

or

$$S_i(\theta_k)/S_i(\theta_{k-1}) < c_l \ \text{ and } \ S_i(\theta_k)/S_i(\theta_{k+1}) < c_l \tag{7}$$

then $S_i(\theta_k)$ is considered an outlier and replaced by $[S_i(\theta_{k-1}) + S_i(\theta_{k+1}0]/2$. Situations described by Eq.(6) are more typical and usually correspond to external interference or transient operational conditions, while those described by Eq.(7) are often just plain measurement errors. Upper and lower threshold values (c_h and c_l, respectively) are adjusted experimentally and depend on the object under consideration. In the following, c_h = 1.5 and c_l = 0.7 were assumed; these values have been found reasonable for steam turbines [10].

As already mentioned, time histories $h(\theta)$ are

determined employing the time window procedure. For a meaningful determination of statistical parameters within such window, however, at least weak stationarity is required. It may be noted that the above definition of an outlier also implicitly implies stationarity. This condition is certainly not fulfilled for θ close to time to breakdown θ_b, as $\theta \to \theta_b \Rightarrow S_i(\theta) \to \infty$. In such circumstances, parameters such as mean value and standard deviation no longer describe statistical properties of a random variable. For this reason, S_i has to be replaced by the trend-normalized S_i', according to [8]

$$S_i'(\theta) = S_i(\theta)\frac{S_{it}(0)}{S_{it}(\theta)} \quad , \tag{8}$$

where lower index t indicates values determined from monotonic trend. $S_{it}(\theta)$ is determined by fitting a monotonically increasing function to experimental values of S_i.

Apart from peak trimming and trend normalization, experimental time histories $S_i(\theta)$ are normalized with respect to their initial values $S_i(0)$. This allows for comparing symptoms of different physical origins, as all normalized symptoms are dimensionless. Normalized symptoms are indicated by lower-case symbols. In order to avoid an influence of a possible outlier on $S_i(0)$, its value is determined as a mean of first three measurements. Fig. 2 shows the results of above-described pre-processing of measurement data.

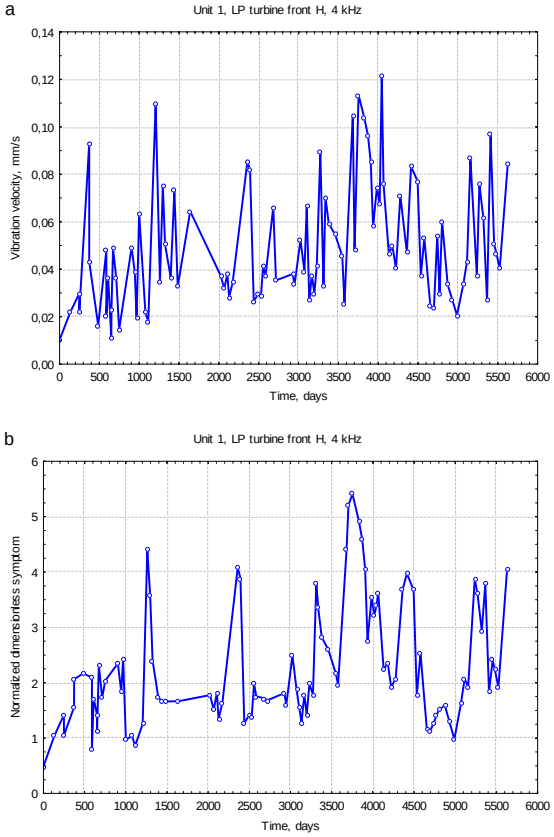

Fig. 2. Effect of peak trimming and trend normalization: raw (a) and pre-processed (b) data. Exponential function $s_{it}(\theta)$ has been used

3. REPRESENTATIVENESS FACTOR

In a descriptive manner we may refer to the entropy decrease with time as an organization of symptom time history around some monotonically increasing curve with a vertical asymptote. Degree of this organization, of which entropy is a measure, increases with lifetime consumption. However, increase rate of the symptom itself is also of importance. Organization may take place around a curve that is only weakly increasing. Such symptom, comparatively insensitive to object condition evolution, would have been of little use. An example is shown in Fig.3a. It is easily seen that there is some entropy decrease, but at the same time symptom value fluctuates about some almost constant value. In fact this symptom reveals weak only very weak increasing trend (see Fig.3b).

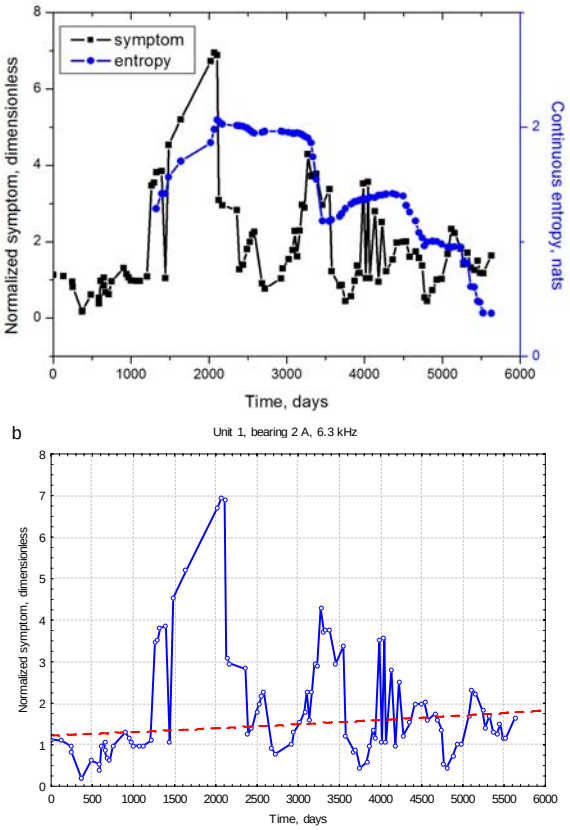

Fig. 3. (a) Plot of normalized symptom and entropy vs. time, showing entropy decrease and low symptom sensitivity to object condition; (b) exponential approximation of normalized symptom vs. time (broken line)

A more elaborate index is thus necessary that would combine information content and symptom increase rate measures. Such suggestion was put forward by the author in a previous study [8]. Initially exponential approximation was used for symptoms which, however, fails as $\theta \to \theta_b$. It is therefore proposed to use linear approximation for entropy:

$$h \approx h(0) - A \cdot \theta \quad (A > 0) \quad , \tag{9}$$

Weibull approximation for normalized symptom:

$$s(\theta) \approx [\ln(1/(1-\theta/\theta_b)]^{1/\gamma} \qquad (10)$$

and define representativeness factor as $R = A/\gamma$. Obviously R should be positive (excluding situations where *both* A and γ are negative); the larger R, the more representative is the symptom under consideration.

4. EXAMPLE

4.1. Object and data acquisition

The suitability of the above-described approach was tested with data obtained for steam turbine fluid-flow system. The object under consideration was a 230 MW condensing unit, operated by a large utility power plant. It comprises high-pressure (HP), intermediate-pressure (IP) and low-pressure (LP) turbines, with shaft supported by seven journal bearings. Photo of the turbine-generator unit is shown in Fig.4.

Fig. 4. Photo of the 230 MW turbine-generator unit (source: *Wikipedia*)

The unit was commissioned in 1998 and acquisition of vibration data started soon afterwards. Vibration velocity spectra were recorded at points located on turbine bearings and LP casing (front and rear part). Frequency range was set at 10 kHz and 23% CPB spectra were used. Available database covers a period of over sixteen years, with average time interval between successive measurements of about 56 days.

In the following, attention shall be focused on the LP turbine. Due to lower temperature and comparatively low pressure gradients, deterioration of the LP fluid-flow system technical condition is slower than for HP and IP ones. Moreover, much weaker influence of control should be expected [11]. In practice vibration time histories recorded at points associated with the LP turbine in the blade frequency range are somehow more regular than for HP and IP turbines and the period of sixteen years should be long enough to detect an increasing trend. It has to be noted that LP casing was not opened during the period under consideration, so it may be assumed that we are dealing with a single life cycle. This is important, as suitable normalization should be otherwise performed [3].

According to the turbine vibrodiagnostic model [12], LP fluid-flow system generates components that are contained in five 23% CPB bands, with mid-frequencies of 1.6 kHz, 2.5 kHz, 3.15 kHz, 4 kHz and 6.3 kHz. Given four measurement points (front LP bearing, LP casing front/rear and rear LP bearing) and three directions (vertical, horizontal and axial), we arrive at sixty distinct symptoms.

4.2. Preliminary selection of symptoms

For the sake of clarity it was decided to perform a preliminary selection of symptoms, employing the SVD method. Details of relevant procedures may be found in author's previous papers (see e.g. [8]). It has been found that the contributions of first three singular values into generalized damage amount to 28.5, 10.5 and 6 percent, respectively. Combined contributions of first six singular values are about 60 percent. This suggests that the dominating damage mechanism has already appeared. On this basis, twelve symptoms with the highest contributions into first three singular values have been determined. They are listed in Table 1.

Table 1. Results of preliminary symptoms selection

Symptom No.	Point	Direction	Mid-freq. [kHz]
1	Front bearing	vertical	6.3
2		horizontal	4.0
3		horizontal	6.3
4		axial	6.3
5	Casing front	vertical	4.0
6		vertical	6.3
7		horizontal	4.0
8		axial	4.0
9	Casing rear	horizontal	4.0
10		axial	4.0
11	Rear bearing	vertical	6.3
12		horizontal	4.0

It is noteworthy that all twelve symptoms selected in this manner represent the 4 kHz and 6.3 kHz frequency bands. This implies that fluid-flow system degradation processes are more pronounced for first two LP turbine stages [12].

The following ICM analysis has been performed for twelve symptoms listed in Table 1.

4.3. ICM analysis

Calculations of differential entropy for all selected symptoms have been performed with the following assumptions:

- time window length: 25 consecutive data points;
- distribution type: normal;
- peak trimming thresholds: $c_h = 1.5$, $c_l = 0.7$;
- trend normalization: exponential;
- logarithm base: e (entropy given in nats).

Results are presented graphically in Fig.5. As one common drawing for all twelve symptoms would have been rather difficult to interpret, it has been divided into four parts, each for three symptoms.

Cursory and qualitative examination of entropy time histories shown allows for distinguishing three types of behaviour:

- erratic (comparatively large fluctuations, but no marked increasing or decreasing trend: symptoms Nos. 3, 4 and 6);
- weakly time-dependent (small and slow variations, lack of pronounced increasing or decreasing trend: symptoms Nos. 7 and 10);
- decreasing (with various decrease rates: symptoms Nos. 1, 2, 5, 8, 9, 11 and 12).

Moreover, it can be seen that in some cases there is a slight entropy increase starting at about $\theta = 4500$ days. The reasons of this phenomenon are unclear. As already mentioned, LP turbine casing was not opened during the entire period under consideration. On this basis it was assumed that we are dealing with a single life cycle and thus a continuous smooth symptom life curve. However, other activities cannot be excluded that influence vibration spectra in the blade frequency range recorded in accessible points. No traces of such activities have been found. In general, such occurrence should result in an abrupt, stepwise change of $S_i(\theta)$. It has already been suggested [8] that some method for detecting such changes could be applied. One method of this type, known as CUSUM (Cumulative Sum Control Chart) (see e.g.[13]) is currently studied by the author.

Bearing in mind that the linear approximation of entropy given by Eq.(9) may be a rough one, values of the coefficient A for above-mentioned twelve symptoms have been calculated and given in Table 2.

Table 2. Linear approximation coefficient values

Symptom No.	A ($\times 10^{-4}$) [nat/day]
1	4.321
2	3.290
3	-0.874
4	1.506
5	0.603
6	1.345
7	0.054
8	1.183
9	2.578
10	-0.031
11	2.090
12	0.368

From Table 2 it is easily seen that values of the linear approximation coefficient vary within quite broad limits. Symptoms Nos. 3 and 10 can be immediately excluded, as A is for them negative (albeit small). Best results have been obtained for symptoms Nos. 1, 2, 9 and 11.

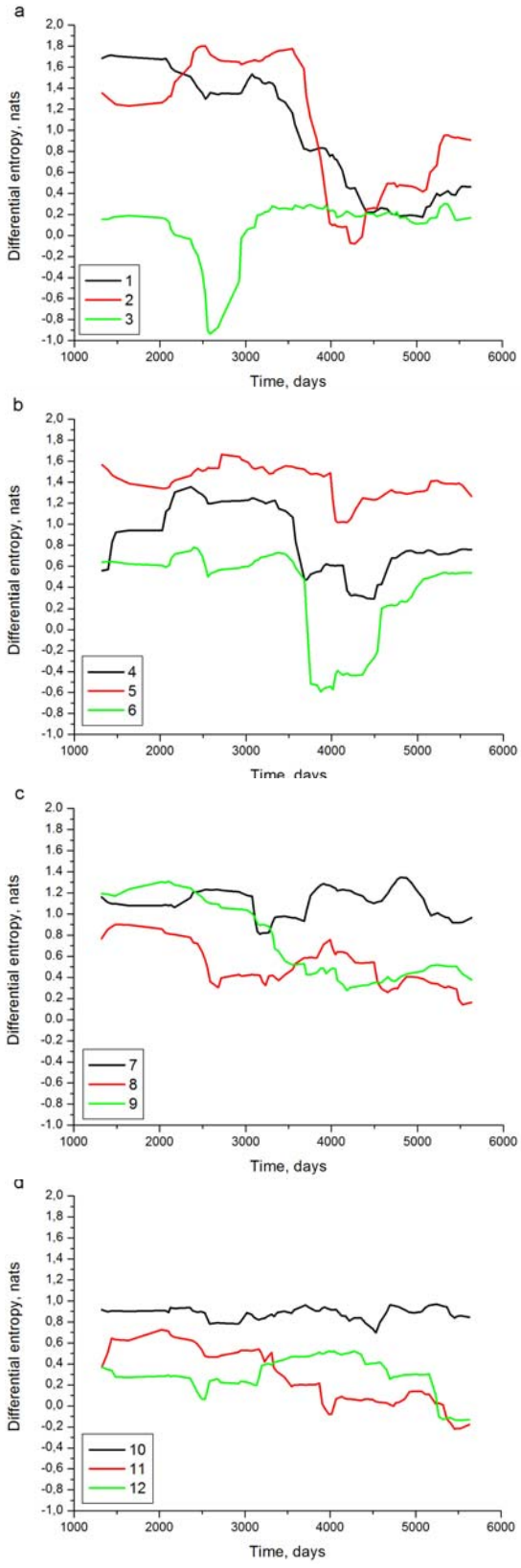

Fig. 5. Differential entropy time histories for symptoms listed in Table 1; (a) symptoms 1 – 3; (b) symptoms 4 – 6; (c) symptoms 7 – 9; (d) symptoms 10 – 12

Table 3 lists representativeness factor values, calculated with the assumption of Weibull symptom life curve. Four values are negative, which means

that relevant symptoms should be excluded. In fact this is the case with symptom No. 2, which has comparatively high entropy decrease rate. The highest value is obtained for symptom No. 1, with symptom No. 11 ranking second. These should be considered most representative for lifetime consumption assessment.

Table 3. Representativeness factor values

Symptom No.	R ($\times 10^{-4}$) [nat/day]
1	5.490
2	-0.563
3	-0.945
4	1.540
5	0.931
6	1.306
7	0.025
8	0.479
9	1.555
10	-0.001
11	2.761
12	-0.012

3. SUMMARY

The ICM method is comparatively simple and yields reasonable results. On the other hands, there are several things that might be improved. It may be noted that entropy time histories are in many cases rather irregular. One possible reason has already been discussed. Another one may be related to the time window procedure. The window containing only 25 data points is rather small, but available database does not seem to allow for a much broader one. Data pre-processing, in particular removing of outliers, might also be modified (see e.g. [14]). With this in mind, the concept seems interesting and deserving further studies.

REFERENCES

[1] Cempel C.: Multidimensional condition monitoring of mechanical systems in operation. *Mechanical Systems and Signal Processing*, vol. 17, No. 6, 2003, pp. 1291-1303.
[2] Gałka T.: Application of the Singular Value Decomposition method in steam turbine diagnostics. *Proceedings of the CM2010/ MFPT2010 Conference*, Stratford-upon-Avon, UK, 2010, paper No. 107.
[3] Gałka T.: *Evolution of Symptoms in Vibration-Based Turbomachinery Diagnostics*. ITeE Publishing House, Radom, 2013.
[4] Rényi A.: On measures of entropy and information. *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, 1961, pp. 547-561.
[5] Shannon C.E., Weaver W.: *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, USA, 1949.
[6] Maszczyk T., Duch W.: Comparison of Shannon, Rényi and Tsallis entropies used in decision trees, *Lecture Notes in Computer Science*, vol. 5097, 2008, pp. 643-651
[7] Taneja. I.J.: *Generalized Information Measures and Their Applications*, on-line book: www.mtm.ufsc.br/~taneja/book/book.html, 2001
[8] Gałka T.: A comparison of two symptom selection methods in vibration-based turbomachinery diagnostics. *Journal of Vibroengineering*, vol. 17, issue 7, 2015, pp. 3505-3514.
[9] Hodge V.I., Austin J.: A survey of outlier detection methodologies, *Artificial Intelligence Review*, vol. 22, 2004, pp. 85-126
[10] Gałka T: On the application of Shannon entropy and continuous entropy in the evaluation of diagnostic symptoms. *International Journal of Condition Monitoring*, vol. 5, No. 3, October 2015, pp. 12-17
[11] Gałka T.: Influence of load and interference in vibration-based diagnostics of rotating machines. *Advances and Applications in Mechanical Engineering and Technology*, vol. 3, No. 1, 2011, pp. 1-19
[12] Orłowski Z.: A model for operational diagnostics of steam turbines. *Mechanical Systems and Signal Processing*, vol. 9, 1995, pp. 215-222
[13] *NIST/SEMATECH e-Handbook of Statistical Methods*, www.itl.nist.gov/div898/handbook/, January 21, 2016
[14] Maronna R.A., Martin D., Yohai V.J.: *Robust Statistics*. *Theory and Methods*, John Wiley, 2006

Received 2016-04-28
Accepted 2016-09-02
Available online 2016-09-19

Tomasz GAŁKA is an associate professor and Director of the Institute of Power Engineering in Warsaw. His entire professional career has been associated with diagnostics of rotating machines, mainly large steam turbines. His main interests include vibration-based diagnostics, lifetime consumption assessment and statistical analysis of machine condition symptoms.